$Y = X\beta^* + \varepsilon$, $\varepsilon \sim N(0,\sigma^2 I_n)$

$\mathbb{R}^n$ $\mathbb{R}^{n\times p}$ $\mathbb{R}^p$ $\mathbb{R}^n$, β^*, σ^2 unknown

Given: $u \in \mathbb{R}^p$, assume X has full column rank

Goal: Find non-asymptotic hypothesis test

$H_0: u^T\beta^* \leq 0$ vs $H_1: u^T\beta^* > 0$

① Find estimators for β^*, σ^2

(2a) Linear algebra review

(2b) Find pivot, T_n

$\psi = \{T_n > s\}$

③ Adjust s to match level α

Hypothesis test in Linear Regression

$(X_1, Y_1), \dots, (X_n, Y_n)$; $X_i \in \mathbb{R}^p$, $Y_i \in \mathbb{R}$

$Y_i = X_i^T\beta^* + \varepsilon_i$, $\beta^* \in \mathbb{R}^p$, $\varepsilon_i \sim N(0,\sigma^2)$ iid.

$$\underbrace{\underset{\mathbb{R}^n}{\underset{\cap}{\begin{pmatrix}Y_1\\ \vdots\\ Y_n\end{pmatrix}}}}_{=Y} = \begin{pmatrix}X_1^T\beta^*\\ \vdots\\ X_n^T\beta^*\end{pmatrix} + \begin{pmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{pmatrix} = \underbrace{\begin{pmatrix}X_1^T\\ \vdots\\ X_n^T\end{pmatrix}}_{X \in \mathbb{R}^{n\times p}} \underset{\mathbb{R}^p}{\underset{\cap}{\beta^*}} + \underbrace{\begin{pmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{pmatrix}}_{=\varepsilon \in \mathbb{R}^n}$$

$\varepsilon \sim N(0, \sigma^2 \cdot I_n)$

Given $u \in \mathbb{R}^p$, test $H_0: u^T\beta^* \leq 0$ $\quad H_1: u^T\beta^* > 0$

e.g. $u = e_1 = \begin{pmatrix}1\\ 0\\ \vdots\\ 0\end{pmatrix} \in \mathbb{R}^p$: $\quad \beta_1^* \leq 0 \quad \beta_1^* > 0$

Assume

X full column rank

$$\underset{\mathbb{R}^n}{Y} = \underset{\mathbb{R}^{n\times p}}{X}\underset{\mathbb{R}^p}{\beta^*} + \underset{\mathbb{R}^n}{\varepsilon},\ \varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$$

β^*, σ^2 unknown, given $\mu \in \mathbb{R}^p$

<u>Goal</u>: Test $H_0 : \mu^T\beta^* \leq 0$

$H_1 : \mu^T\beta^* > 0$

① Estimators for β^*, σ^2

$$\hat\beta = (X^TX)^{-1}X^TY$$

$$\hat\sigma^2_{mle} = \frac{1}{n}\|Y - X\beta\|_2^2$$

(2a) Linear algebra review

(2b) Pivot, test statistic T_n

$$\psi = \{T_n > s\}$$

③ Adjust s

① Maximum likelihood estimators

$$\mathbb{P}_{\beta,\sigma^2}(Y_1, \dots, Y_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\cdot(Y_i - X_i^T\beta)^2}$$

$$\ell_{\beta,\sigma^2}(Y_1,\dots,Y_n) = \sum_{i=1}^{n}\left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\cdot(Y_i - X_i^T\beta)^2\right]$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\|Y - X\beta\|_2^2$$

$$\nabla_\beta \ell_{\beta,\sigma^2} = +\frac{2}{2\sigma^2}X^T(Y - X\beta) \overset{!}{=} 0 \Rightarrow X^TY - X^TX\hat\beta = 0$$

$$\Rightarrow X^TX\hat\beta = X^TY$$

$$\overset{X \text{ full col rk}}{\Rightarrow} \boxed{\hat\beta = (X^TX)^{-1}X^TY}$$

$$\partial_{\sigma^2}\ell_{\beta,\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\|Y - Y\beta\|_2^2 \overset{!}{=} 0 \Rightarrow \boxed{\hat\sigma^2_{mle} = \frac{1}{n}\|Y - X\hat\beta\|_2^2}$$

$$\underset{\mathbb{R}^n}{Y} = \underset{\mathbb{R}^{n\times p}}{X}\underset{\mathbb{R}^p}{\beta^*} + \underset{\mathbb{R}^n}{\varepsilon},\ \varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$$

β^*, σ^2 unknown, given $u \in \mathbb{R}^p$

Goal: Test $H_0: u^T\beta^* \le 0$

$H_1: u^T\beta^* > 0$

① Estimators for β^*, σ^2

$$\hat\beta = (X^TX)^{-1}X^TY$$

$$\hat\sigma^2_{mle} = \frac{1}{n}\|Y - X\hat\beta\|_2^2$$

②a Linear algebra review

②b Pivot, test statistic T_n

$\psi = \{T_n > s\}$

③ Adjust s

②a Linear algebra review

$v_1, \dots, v_k \in \mathbb{R}^m$ orthogonal iff $v_i^T v_j = 0,\ i \neq j$

orthonormal iff — " — & $v_i^T v_i = \|v_i\|_2^2 = 1$

$$V = [v_1, \dots, v_k],\ \|V \underset{\mathbb{R}^k}{\underset{\uparrow}{x}}\|_2^2 = x^T \underbrace{V^T V}_{= I_k} x = x^T x = \|x\|_2^2$$

$$V = [v_1, \dots, v_m] \Rightarrow V \text{ orthogonal} \quad V^T V = I_m = V V^T$$

Spectral theorem: $A \in \mathbb{R}^{m \times m}$, $A^T = A \Rightarrow \exists V$ orthogonal, Λ diagonal

$$\text{s.t. } A = V \Lambda V^T \Leftrightarrow \underbrace{AV}_{= [Av_1 \dots Av_m]} = \underbrace{V \Lambda}_{[\lambda_1 v_1 \dots \lambda_m v_m]}$$

$$A v_i = \lambda_i v_i$$

$Y = X\beta^* + \varepsilon, \varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$

$\mathbb{R}^n \quad \mathbb{R}^{n\times p} \quad \mathbb{R}^p \quad \mathbb{R}^n$

β^*, σ^2 unknown, given $\mu \in \mathbb{R}^p$

Goal: Test $H_0: \mu^T\beta^* \le 0$

$H_1: \mu^T\beta^* > 0$

① Estimators for β^*, σ^2

$\hat{\beta} = (X^TX)^{-1}X^TY$

$\hat{\sigma}^2_{mle} = \frac{1}{n}\|Y - X\beta\|_2^2$

②a Linear algebra review

②b Pivot, test statistic T_n

$\psi = \{T_n > s\}$

③ Adjust s

②a Linear algebra review

Projection matrices $H \in \mathbb{R}^{m\times m}$ proj. matrix iff $H^T = H$, $H^2 = H$

Then, all eigenvalues λ_i of H are either +1 or 0.

⌐Pick eigenvector v_i of H

$$\lambda_i v_i v_i^T = v_i^T H v_i = v_i^T H^2 v_i = v_i^T H^T H v_i = (Hv_i)^T (Hv_i)$$

$$\| \quad \lambda_i \cdot \|v_i\|_2^2$$

$$= (\lambda_i v_i)^T \cdot (\lambda_i v_i)$$

$$= \lambda_i^2 \cdot \|v_i\|_2^2$$

$$\Rightarrow \lambda_i = \lambda_i^2 \Rightarrow \lambda_i \in \{0, 1\}$$

Trace: $tr(A) = \sum_{i=1}^{m} A_{i,i}$, $\quad A \in \mathbb{R}^{m\times m}$

A, B, C

$\mathbb{R}^{n\times m} \quad \mathbb{R}^{m\times l} \quad \mathbb{R}^{l\times n}$

$$\Rightarrow tr(ABC) = tr(CAB) = tr(BCA$$

$$A = V\Lambda V^T, \quad tr(A) = tr(V\Lambda V^T) = tr(V^T V \Lambda)$$

$$= tr(\Lambda) = \sum_{i=1}^{m} \lambda_i$$

$\underset{\mathbb{R}^n}{Y} = \underset{\mathbb{R}^{n\times p}}{X}\underset{\mathbb{R}^p}{\beta^*} + \underset{\mathbb{R}^n}{\varepsilon}, \ \varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$

β^*, σ^2 unknown, given $\mu \in \mathbb{R}^p$

Goal: Test $H_0: \mu^T\beta^* \leq 0$

$H_1: \mu^T\beta^* > 0$

① Estimators for β^*, σ^2

$\hat{\beta} = (X^TX)^{-1}X^TY$

$\hat{\sigma}^2_{mle} = \frac{1}{n}\|Y - X\hat{\beta}\|_2^2$

②a Linear algebra review

②b Pivot, test statistic T_n

$\psi = \{T_n > s\}$

③ Adjust s

②b $\mu^T\hat{\beta} = \mu^T(X^TX)^{-1}X^T(X\beta^* + \varepsilon) = \mu^T\beta^* + \mu^T(X^TX)^{-1}X^T\varepsilon$

$\Rightarrow \mu^T\hat{\beta} - \mu^T\beta^* \sim N(0, \sigma^2\mu^T(X^TX)^{-1}X^TX(X^TX)^{-1}\mu) = \mathcal{N}(0, \sigma^2\mu^T(X^TX)^{-1}\mu)$

$\|Y - X\hat{\beta}\|_2^2 = \|X\beta^* + \varepsilon - X(X^TX)^{-1}X^T(X\beta^* + \varepsilon)\|_2^2$

$= \|\varepsilon - \underbrace{X(X^TX)^{-1}X^T}_{=H}\varepsilon\|_2^2 = \|(I_n - H)\varepsilon\|_2^2$

H is symmetric: $H^T = X\cdot(X^TX)^{-1}X^T = H$,

projection: $H^2 = X\underbrace{(X^TX)^{-1}X^TX}_{=I_p}(X^TX)^{-1}X^T = H$

By spectral thm, $\exists V$ orthogonal, Λ diagonal: $H = V\Lambda V^T = V\begin{bmatrix} \underbrace{1 \ \ddots \ 1}_{\ell} & 0 \\ 0 & 0 \ \ddots \ 0 \end{bmatrix}V^T$

$\ell = tr(H) = tr(X(X^TX)^{-1}X^T) = tr(X^TX\cdot(X^TX)^{-1}) = tr(I_p) = p$

$H = [\underbrace{V_1}_{T} \ \underbrace{V_2}_{n-p}]\cdot\begin{bmatrix} I_p & 0 \\ 0 & 0_{n-p\times n-p}\end{bmatrix}\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = V_1V_1^T$

$\underset{\mathbb{R}^n}{Y} = \underset{\mathbb{R}^{n\times p}}{X}\underset{\mathbb{R}^p}{\beta^*} + \underset{\mathbb{R}^n}{\varepsilon},\ \varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$

β^*, σ^2 unknown, given $\mu \in \mathbb{R}^p$

Goal: Test $H_0: \mu^T\beta^* \le 0$

$H_1: \mu^T\beta^* > 0$

① Estimators for β^*, σ^2

$\hat\beta = (X^TX)^{-1}X^TY$

$\hat\sigma^2_{ub} = \frac{1}{(n-p)}\|Y - X\hat\beta\|_2^2$

(2a) Linear algebra review

(2b) Pivot, test statistic T_n

$\psi = \{T_n > s\}$

③ Adjust s

(2b) $\|Y - X\hat\beta\|_2^2 = \|(I_n - H)\varepsilon\|_2^2$, $H = [V_1\ V_2]\underbrace{\begin{bmatrix} I_p & \\ & 0_{n-p}\end{bmatrix}}_{\Lambda}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix} = V_1V_1^T$

$$\Rightarrow I_n - H = I_n - V\Lambda V^T = VV^T - V\Lambda V^T$$

$$= V[I_n - \Lambda]V^T = [V_1\ V_2]\begin{bmatrix}0_p & \\ & I_{n-p}\end{bmatrix}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix}$$

$$= V_2V_2^T$$

$$\Rightarrow \|Y - X\hat\beta\|_2^2 = \|\underset{n\times(n-p)}{V_2}V_2^T\varepsilon\|_2^2 = \|V_2^T\varepsilon\|_2^2,\ \mathbb{E}[V_2^T\varepsilon] = 0$$

$$\mathrm{Cov}(V_2^T\varepsilon) = \mathbb{E}[V_2^T\varepsilon\varepsilon^TV_2] = V_2^T\mathrm{Cov}(\varepsilon)V_2 = \sigma^2 V_2^TV_2 = \sigma^2 I_{n-p}$$

$$\Rightarrow \frac{1}{(n-p)\sigma^2}\|Y - X\hat\beta\|_2^2 \sim \chi^2_{n-p} \Rightarrow \frac{n}{(n-p)\sigma^2}\hat\sigma^2_{mle} \sim \chi^2_{n-p}$$

$$\mu^T\hat\beta = \mu^T(X^TX)^{-1}X^T(X\beta^* + \varepsilon) = \mu^T\beta^* + (X^TX)^{-1}X^T\varepsilon$$

$$= \mu^T\beta^* + (X^TX)^{-1}X^TX(X^TX)^{-1}X^T\varepsilon = \mu^T + (X^TX)^{-1}X^TH\varepsilon$$

$H\varepsilon \perp\!\!\!\perp V_2^T\varepsilon$

$$\mathrm{Cov}(H\varepsilon, V_2^T\varepsilon) = \mathbb{E}[H\varepsilon\varepsilon^TV_2] = \mathbb{E}[V_1V_1^T\underbrace{\varepsilon\varepsilon^T}_{\mathbb{E}[\varepsilon\varepsilon^T] = \sigma^2 I}V_2] = 0 \Rightarrow$$

$$\Rightarrow \boxed{\hat\sigma^2_{ub} \perp \hat\beta}$$

$\underset{\mathbb{R}^n}{Y} = \underset{\mathbb{R}^{n\times p}}{X}\underset{\mathbb{R}^p}{\beta^*} + \underset{\mathbb{R}^n}{\varepsilon}, \varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$

β^*, σ^2 unknown, given $\mu \in \mathbb{R}^p$

Goal: Test $H_0: \mu^T\beta^* \leq 0$

$H_1: \mu^T\beta^* > 0$

① Estimators for β^*, σ^2

$\hat\beta = (X^TX)^{-1}X^TY$

$\hat\sigma^2_{ub} = \frac{1}{(n-p)}\|Y - X\hat\beta\|_2^2$

(2a) Linear algebra review

(2b) Pivot, test statistic T_n

$T_n = \frac{\mu^T\hat\beta}{\hat\sigma_{ub}\sqrt{\mu^T(X^TX)^{-1}\mu}}$

$\psi = \{T_n > s\}$

③ Adjust s

$s = q_\alpha(t_{n-p})$.

(2b) $\frac{\hat\sigma^2_{ub}}{\sigma^2} \sim \chi^2_{n-p}$, $\hat\sigma^2_{ub} \perp\!\!\!\perp \hat\beta$

$\mu^T\hat\beta - \mu^T\beta^* \sim \mathcal{N}(0, \sigma^2\mu^T(X^TX)^{-1}\mu)$

$\Rightarrow \frac{\mu^T\hat\beta - \mu^T\beta^*}{\sigma\sqrt{\mu^T(X^TX)^{-1}\mu}} \sim \mathcal{N}(0,1) \Rightarrow \frac{\mu^T\hat\beta - \mu^T\beta^*}{\hat\sigma_{ub}\cdot\sqrt{\mu^T(X^TX)^{-1}\mu}} \sim t_{n-p}$

$T_n = \frac{\mu^T\hat\beta}{\hat\sigma_{ub}\cdot\sqrt{\mu^T(X^TX)^{-1}\mu}}$

③ Case 1: $\mu^T\beta^* = 0$: $\mathbb{P}(T_n > s) = \mathbb{P}\left(\frac{\mu^T\hat\beta - \mu^T\beta^*}{\hat\sigma_{ub}\sqrt{\mu^T(X^TX)^{-1}\mu}} > s\right)$

$= \mathbb{P}(Z > s), Z \sim t_{n-p}, \boxed{s = q_\alpha(t_{n-p})}$, $1-\alpha$ quantile of t_{n-p}.

$\overset{!}{=} \alpha$

Case 2: $\mu^T\beta^* < 0$: $\mathbb{P}(T_n > s) = \mathbb{P}\left(\frac{\mu^T\hat\beta}{\hat\sigma_{ub}\sqrt{\mu^T(X^TX)^{-1}\mu}} > s\right)$

$\leq \mathbb{P}\left(\frac{\mu^T\hat\beta - \mu^T\beta^*}{\hat\sigma_{ub}\sqrt{\mu^T(X^TX)^{-1}\mu}} > s\right) = \alpha$

$H_0: \beta_1 \leq \beta_2$ vs. $H_1: \beta_1 > \beta_2$